IMPORTANT: The 2-step dermoscopy algorithm aims to guide the decision-making process to biopsy or not biopsy a skin lesion by providing the most probable diagnosis via a systematic approach.

OBJECTIVE: To evaluate the diagnostic accuracy and potential limitations of the first step (to differentiate melanocytic from nonmelanocytic lesions) of the 2-step dermoscopy algorithm.

DESIGN, SETTING, AND PARTICIPANTS: Retrospective study in a clinical practice of one dermatologist of biopsy data of all skin lesions from one clinic during a 10-year period. The prebiopsy and histopathology diagnoses were classified as melanocytic or nonmelanocytic.

MAIN OUTCOMES AND MEASURES: The diagnostic accuracy, sensitivity, specificity, positive predictive value, and negative predictive value for the first step were estimated using the histopathological lesion classifications as the standard.

RESULTS: The sensitivity of the first step for correctly identifying melanocytic lesions was 85%, and the specificity was 94%. Approximately 7% of all lesions (667 of 9168) had discordant classifications, with 415 (4.5%) being false-positive lesions (clinically classified as melanocytic and histopathologically classified as nonmelanocytic) and 252 (2.7%) being false negatives (clinically classified as nonmelanocytic and histopathologically classified as melanocytic). Common classification errors included intradermal nevus misclassified as basal cell carcinoma and nonmelanocytic lesions (eg, seborrheic keratosis, lichen planus-like keratosis, basal cell carcinomas) misclassified as melanocytic because they mimic melanoma. Clinically, 8 of 381 melanomas were misclassified as nonmelanocytic (primarily as pigmented basal cell carcinomas and squamous cell carcinomas).

CONCLUSIONS AND RELEVANCE: The 2-step dermoscopy algorithm, including its first step, has high sensitivity, specificity, and accuracy and can be relied on to provide an accurate and specific prebiopsy diagnosis and to help guide management decisions. Some lesions had a higher chance of being misclassified, with the most common being intradermal nevi. This algorithm helps toward maximizing the detection of skin cancer to ensure that malignant lesions are not missed and aims at making more precise clinical diagnoses.

Author Affiliations: Department of Dermatology and Cutaneous Surgery, University of Miami, Miami, Florida (Chen); Dermatology Service, Memorial Sloan Kettering Cancer Center, New York, New York (Dusza, Marghoob); Dermatology Service, Aurora Skin Cancer Center, Universidad Pontificia Bolivariana, Medellin, Colombia (Jaimes).

Corresponding Author: Ashfaq A. Marghoob, MD, Dermatology Service, Memorial Sloan Kettering Cancer Center, 16 E 60th St, Third Floor, New York, NY 10022 (marghooa@mskcc.org).
Dermoscopy is a noninvasive technique that enables clinicians to evaluate skin lesions and improve their diagnostic accuracy, compared with naked-eye clinical examination. The 2-step dermoscopy algorithm is one dermoscopic method intended to guide the observer toward the most likely diagnosis and to help in the biopsy decision-making process. This algorithm was introduced in 2001 by a panel of the virtual Consensus Net Meeting on Dermoscopy, and it has since undergone minor modifications. While its primary aim is to help clinicians avoid missing the diagnosis of melanoma, important secondary aims are to segregate lesions into melanocytic and nonmelanocytic categories and to provide the observer with the most probable diagnosis via the use of a systematic approach that is easy to implement.

The goal of the first step of the 2-step dermoscopy algorithm is to help differentiate melanocytic from nonmelanocytic skin lesions. An 8-level criteria ladder was created to assist observers in classifying lesions as melanocytic or nonmelanocytic. The nonmelanocytic benign lesions (ie, dermatofibromas, seborrheic keratoses [SK], hemangiomas, and clear cell acanthomas) and cancers (basal cell carcinoma [BCC] and squamous cell carcinoma [SCC]) are diagnosed as such in the first step. Lesions that are classified as melanocytic are further evaluated in the second step, which is intended to differentiate benign melanocytic neoplasms from melanoma. There are multiple methods that can be used to differentiate nevi from melanoma, including pattern analysis, a 7-point checklist, and the method by Menzies et al, just to mention a few. While these various algorithms that can be used in the second step of the 2-step algorithm have been extensively studied and shown to be fairly accurate, the accuracy of the first step (to differentiate melanocytic from nonmelanocytic lesions) has been evaluated in only one study, which demonstrated a high sensitivity but a low specificity. A major critique of the 2-step algorithm is that initial errors in the classification of melanocytic status will render the second step irrelevant. The primary objectives of this study were to evaluate the diagnostic accuracy of the first step of the 2-step algorithm and to highlight the potential limitations of the first step using data from 10 years of dermoscopy-guided clinical practice.

Methods

This study was performed in accord with the Institutional Review Board at Memorial Sloan Kettering Cancer Center, and a retrospective study waiver for informed consent was granted. All skin lesions evaluated via the 2-step dermoscopy algorithm and biopsied in the clinic of one dermatologist (A.A.M.) were retrospectively identified between January 1, 2001, and December 31, 2010, at a high-risk dermatology clinic at Memorial Sloan Kettering Cancer Center. This dermatologist has been using the 2-step algorithm in his daily practice during that time frame and has relied on it for rendering a prebiopsy clinical diagnosis for each biopsied lesion. The prebiopsy clinical diagnosis for all lesions was recorded in the physician's biopsy ledger before biopsy. This prebiopsy diagnosis was subsequently retrospectively compared with the histopathology for that lesion.

The prebiopsy clinical diagnoses were categorized as melanocytic or nonmelanocytic. Lesions were considered melanocytic if any of the dermoscopic structures of melanocytic neoplasms were present (eTable 1 in the Supplement). If a lesion manifested none of these structures and did not manifest the diagnostically features of dermatofibroma, BCC, SCC, SK, hemangioma, or clear cell acanthoma, then blood vessels were evaluated. The morphology, arrangement, and distribution of the vessels within an isolated neoplasm were analyzed. In general, the presence of predominantly arborizing vessels, glomerular vessels, hairpin vessels with a white halo, or crown vessels led to the classification of the lesion as nonmelanocytic. The presence of comma vessels (slightly curved vessels), dotted vessels (red dots of 0.01-0.02 mm), serpentine vessels (linear irregular or undulating short vessels), milky-red globules and vascular blush (ill-defined globules and areas of milky-red), Corkscrew vessels (coiled and tortuous vessels), and polymorphous vessels (combination of ≥2 vessel morphologies) led to the classification of the lesion as melanocytic. Finally, if the lesion did not reveal any of the aforementioned structures, it was considered nonclassifiable, and by default all such lesions were considered suspect for melanoma.

Lesions with overlapping prebiopsy clinical diagnoses that included both melanocytic and nonmelanocytic neoplasms (ie, dysplastic nevus vs SK) were deemed clinically indeterminate and, based on the 2-step algorithm, were classified as melanocytic by default. This default is part of the 2-step algorithm by design and is intended to prevent an observer from missing the diagnosis of melanoma. Similarly, lentigines manifesting a network were also categorized as melanocytic in an effort to avoid missing a lentigo maligna. The histopathological diagnosis of each lesion was obtained from pathology reports, and the lesion was then segregated into melanocytic or nonmelanocytic lesions. Those lesions in which the histopathological diagnosis included both melanocytic and nonmelanocytic components (ie, collision lesion between SK and melanocytic nevus) were classified as melanocytic and called “collision/combination.” Any lesion previously biopsied was excluded.

To appraise the first step as a diagnostic test for identifying a melanocytic lesion, each lesion was defined as either true positive (TP), false positive (FP), true negative (TN), or false negative (FN). True positive was defined as lesions classified as melanocytic by clinical diagnosis and confirmed as melanocytic by histopathological diagnosis. False positive was defined as lesions classified as melanocytic by clinical diagnosis but confirmed as nonmelanocytic by histopathological diagnosis. True negative was defined as lesions classified as nonmelanocytic by clinical diagnosis and confirmed as nonmelanocytic by histopathological diagnosis. False negative was defined as lesions classified as nonmelanocytic by clinical diagnosis but confirmed as melanocytic by histopathological diagnosis. The sensitivity, specificity, positive predictive value, and negative predictive value were computed accordingly. The diagnostic accuracy, defined as the proportion of true melanocytic and nonmelanocytic lesions identified correctly out of the total lesion population, was calculated as follows: (TP + TN) / (TP + FP + TN + FN). All performance measures are presented with their 95% binomial exact CIs. Analyses were performed with statistical software (STATA, version 12.1; StataCorp LP).
Results

In total, 9168 lesions were biopsied during the 10-year period, of which 1641 (17.9%) were biopsy-proven melanocytic and 7527 (82.1%) were biopsy-proven nonmelanocytic based on the final histopathological diagnoses (Table 1). There were 81 collision lesions, including 52 containing at least 1 melanocytic diagnosis and 29 containing only a nonmelanocytic clinical diagnosis. Based on the data in Table 1, the sensitivity of the first step of the 2-step algorithm to identify a true melanocytic lesion was 85% (95%CI, 83%-86%), and the specificity to identify a true nonmelanocytic lesion was 94% (95%CI, 94%-95%). The positive predictive value, or the proportion of clinically suspected melanocytic lesions that were truly melanocytic by histopathology, was 77% (95%CI, 75%-79%). The negative predictive value, or the proportion of clinically suspected nonmelanocytic lesions there were histopathologically confirmed as such, was 97% (95%CI, 96%-97%). The diagnostic accuracy of the first step was 93% (95%CI, 92%-93%). If the clinical or histopathological diagnoses for lentigines were reclassified as nonmelanocytic, the values for the sensitivity, specificity, positive predictive value, and negative predictive value did not differ significantly.

Among 9168 biopsied lesions, 667 (7.3%) had discordant clinical and histopathological diagnoses. These lesions were further categorized as FP and FN lesions.

The FP lesions included 415 (4.5%) that were clinically misclassified as melanocytic and proved to be nonmelanocytic by pathology (Table 1). The leading prebiopsy clinical diagnoses for these misclassified nonmelanocytic lesions were melanoma (32.0%), lentigo maligna (23.0%), and nevus of any type, including dysplastic and Spitz (17.0%). The histopathological diagnoses of these 415 lesions were most commonly SK, actinic keratosis, and lichen planus−like keratosis (LPLK) (Table 2). The FN lesions included 252 (2.7%) that were clinically misclassified as nonmelanocytic but proved to be melanocytic by pathology (Table 1). The 2 most common prebiopsy clinical diagnoses for these misclassified melanocytic lesions were BCC (63.0%) and SK (10.0%). The most common histopathological diagnoses for these clinically misclassified lesions were intradermal nevi (IDN) (61.0%) and other nevi, including dysplastic nevus and otherwise unspecified types of nevi (25.0%) (Table 3). In addition, 8 melanomas (2.1% of all 381 melanomas) were misclassified as nonmelanocytic (eTable 2 in the Supplement). These lesions had various clinical diagnoses, including BCC, LPLK, SK, SCC, and pseudolymphoma.

Discussion

The primary objectives of dermoscopy during skin cancer surveillance are to maximize the early detection of skin cancers and to ensure that malignant lesions, especially melanomas, are not missed, while at the same time minimizing the number of biopsies of benign lesions.10,11 The benefits of dermoscopy are irrefutable and have been reviewed in 3 large meta-analyses.1,12,13
which have all shown an improvement in the diagnostic accuracy with its use. The most recent meta-analysis to date, performed by Vestergaard et al, demonstrated that the diagnostic odds ratio for dermoscopy was 15.6 compared with naked-eye examination. From a practical point of view, dermoscopy leads the clinician toward a more accurate prebiopsy diagnosis and helps in management decisions regarding which lesions require a biopsy and which lesions can be safely monitored.

The 2-step dermoscopy algorithm is one dermoscopic method that resulted from the combined experience of dermatoscopists from around the world. Since its introduction in 2001, it has undergone minor modifications. The goal of the first step of the 2-step algorithm is to differentiate melanocytic from nonmelanocytic skin lesions by looking for the presence or absence of specific features. This segregation of lesions into melanocytic or nonmelanocytic categories acts as a triage mechanism and is what many clinicians (even those not using dermoscopy) perform intuitively on a routine basis. This practice ensures that melanocytic lesions and collision tumors are delineated so as not to miss a melanoma. In addition, to differentiate melanocytic from nonmelanocytic lesions, the first step serves as an aid to correctly subclassify and render a more precise clinical diagnosis of the nonmelanocytic lesions (ie, dermatofibroma, BCC, SCC, SK, hemangioma, and clear cell acanthoma). In fact, the results of the present study showed that the first step of the 2-step algorithm has a high diagnostic accuracy (93%), with a sensitivity of 85% and a specificity of 94%. Similarly, Tschantl et al demonstrated that the first step is very sensitive for melanocytic lesions (97%). However, their study showed a low specificity (34% for Australian patients and 68% for European patients), especially for patients with severely sun-damaged skin. Some reasons that may explain this difference include the evaluation of different populations with a higher prevalence of lentigines (lentigines are often misclassified as melanocytic lesions; therefore, populations with many lentigines can account for a lower reported specificity), as well as their smaller sample size (n = 702) and the interpretation of structures present in lesions. On the other hand, another previous study showed that participants were able to correctly classify 95% of melanocytic lesions and more than 90% of nonmelanocytic lesions with good interobserver agreement (κ = 0.63; 95% CI, 0.62-0.63) using the 2-step algorithm.

The misclassification of melanocytic lesions as nonmelanocytic and nonmelanocytic lesions as melanocytic may lead to the misclassification of some malignant lesions as benign, which in turn may be potentially dangerous for the patient. Therefore, the algorithm attempts to avoid this pitfall by classifying all clinically indeterminate lesions (ie, lesions that have insufficient features to be definitively called melanocytic or nonmelanocytic) as melanocytic. This helps ensure that potential melanomas are not missed, although at the cost of diagnostic accuracy. For example, our results showed that out of 113 clinically indeterminate lesions, more than two-thirds (n = 76) were found to be nonmelanocytic by histopathology, including 3 BCCs and 1 SCC. For such lesions, in which the 2-step algorithm is unable to differentiate melanocytic from nonmelanocytic, other algorithms, such as “chaos and clues,” may be of benefit. The chaos and clues method relies on the fact that most pigmented skin cancers (BCC and melanoma) will be dermoscopically asymmetric and will manifest at least 1 dermoscopically concerning structure. However, it has the pitfall that a malignant lesion manifesting a symmetric morphology or an amelanotic lesion will potentially be missed. With that said, it is likely that most clinicians intuitively amalgamate the chaos and clues method and the 2-step dermoscopy algorithm in some fashion.

The misclassification of melanocytic lesions as nonmelanocytic and nonmelanocytic lesions as melanocytic may also lead to the misclassification of some benign lesions as malignant, which in turn may affect the specificity. However, this is necessary to balance the dual purposes of triage and diagnosis for detecting a suspicious skin lesion. For this reason, the first step has been criticized by some for being complex and unreliable, with classification errors that may lead to lowering of diagnostic precision. Specifically, one of the main critiques of the 2-step dermoscopy algorithm is the misclassification of some solar lentigines or SK as melanocytic lesions. While this error occurred in our study (in 117 of 1650), we believe that the primary intent of the 2-step algorithm is to guide the clinician in the biopsy decision-making process regarding a suspected skin lesion, with a secondary intent of improving the prebiopsy diagnostic precision. Although other algorithms such as the revised pattern analysis are designed to maximize prebiopsy diagnostic precision, they are significantly more time-consuming to use on a routine basis. In essence, both methods can lead to the same management decisions, and the only difference is that the prebiopsy clinical diagnostic precision may be slightly lower with the 2-step dermoscopy algorithm. Ultimately, these outcomes do not affect patient care.

<table>
<thead>
<tr>
<th>Clinical Diagnosis</th>
<th>Histopathological Diagnosis</th>
<th>Intradermal Nevus</th>
<th>Dysplastic Nevus</th>
<th>Nevus, Unspecified</th>
<th>Lentigo</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal cell carcinoma</td>
<td>130</td>
<td>5</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Seborrheic keratosis</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Lichen planus−like keratosis</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1b</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Otherc</td>
<td>12</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>

- The most common single clinical diagnoses are shown (29 collision lesions are not shown).
- Atypical proliferation of melanocytes.
- Includes dermatofibroma, skin tag, and poroma.
To expand on the aforementioned issues, our study showed that the misclassification of lesions in the first step may result in an incorrect clinical diagnosis but that the management outcome remained unchanged. For example, pigmented BCC and SCC may enter the differential diagnosis for melanoma, especially when there are dermoscopic features that are also seen in melanocytic lesions (eg, brown or black globules and irregular vessels). Although performing a biopsy of a BCC or SCC to rule out a diagnosis of melanoma can be considered an error, leading to lowering of the clinical diagnostic precision, the final management decision to biopsy remains unaltered. Similarly, biopsying a melanoma to rule out a diagnosis of BCC or SCC has the same effect. Indeed, we identified 8 melanomas that were clinically misclassified as nonmelanocytic; 6 of them had a differential diagnosis that included BCC or SCC (eTable 2 in the Supplement). Among 8 lesions, BCC was the most common prebiopsy clinical diagnosis (n = 5), and all 5 were described as erythematous plaques or papules (Figure 1A and B). The other lesion that sometimes had clinical morphologic features overlapping with melanoma was LPLK (Figure 1C and D). Finally, the patient with suspected pseudolymphoma had a history of a 1.6-mm melanoma removed from the left abdomen with a negative sentinel lymph node 7 years previously (Figure 1E and F). This patient was seen for an acute visit regarding a lesion that proved to be metastatic melanoma. As described in prior retrospective investigations, some melanomas can be difficult to diagnose because they may simulate nonmelanocytic tumors such as hemangioma, BCC, and SK. We suggest careful evaluation of these entities to look for atypical dermoscopic features (poorly demarcated borders, irregular vessels, and blue-gray pigmentation) and to ask about a history of change.

In the present study, several recurrent classification errors emerged that deserve mention. Distinguishing IDN from BCC represented one of the most common errors of benign and malignant misclassification. In fact, in the group of lesions misclassified as nonmelanocytic, approximately 80% of lesions that had a clinical diagnosis of BCC proved to be an IDN by pathology. By clinical morphology, the appearance of BCCs and IDN can often be strikingly similar, manifesting as pink or skin-colored dome-shaped papules (Figure 2). When irritated, IDN may be mistaken for ulcerated BCCs. According to the first step, many IDN may be mistaken for nonmelanocytic lesions because of the lack of melanocytic structures, and this together
with the presence of arborizing vessels (which are commonly seen in IDN) can lead to the incorrect prebiopsy clinical diagnosis of BCC. With that said, in 10 years of full-time clinical practice, this has occurred only 130 times. Another classification error is the concern regarding melanoma among FP lesions. This should come as no surprise because previous observations have shown that melanoma can mimic many nonmelanocytic entities, including BCC, SK, and LPLK. For example, nonpigmented BCC with irregular vessels and crystalline structures can be confused for an amelanotic melanoma. Pigmented BCC with large blue ovoid nests or fuzzy streaks can also mimic melanoma. On occasion, SK can resemble melanocytic lesions by exhibiting a pseudonetwork or irregular vessels, particularly when they are irritated. Finally, LPLK has a morphology that often overlaps with that of melanoma and superficial BCC. In fact, based on clinical and dermoscopic examination of an LPLK, it is often impossible to rule out melanoma or BCC.

When can errors in the first step of the 2-step dermoscopy algorithm occur? The misclassification of nonmelanocytic lesions as melanocytic and vice versa can occur when lesions do not reveal the criteria used to differentiate melanocytic...
from nonmelanocytic lesions, when lesions have overlapping criteria, when they reveal structures mimicking those normally seen in the other category, or when technical issues (ie, applying too much pressure to the dermoscope, thereby impairing visualizing vessels) preclude the observation of a key structure. Other reasons can include perceptual errors (in which the observer does not see a structure that is in fact present), cognitive errors (in which a dermoscopic structure is identified but the observer lacks the knowledge to interpret it), and search satisfaction, anchoring bias, and confirmation bias (also known as belief bias).

Limitations of our study include the retrospective nature of the study. We were able to evaluate only lesions subjected to biopsy, and this introduces a selection bias for lesions that are more difficult to diagnose clinically and dermoscopically. Naturally, we have no way of knowing the misclassification rate of the first step on lesions that were not biopsied. Another limitation is that these results represent the practice of one observer (A.A.M.). While the observer used dermoscopy on every lesion and the 2-step method was a well-integrated part of the analysis of all lesions since the start of the study, his self-learning and adjustments to a certain patient population could have affected the accuracy of the diagnostic algorithm. One must also keep in mind that dermoscopy consists of only one component of the physical examination. Other factors such as a patient's history, clinical context, and palpation of the lesion could have biased the dermoscopic interpretation of some lesions.

Conclusions

In conclusion, the 2-step dermoscopy algorithm was primarily designed to ensure that malignant lesions are not missed and, second, to aid in making a more precise clinical diagnosis. In essence, the algorithm helps toward maximizing the detection of skin cancer, while lowering the number of biopsies of benign lesions. Our results demonstrate that the first step of the 2-step dermoscopy algorithm is highly sensitive and specific and support its continued use. A few pitfall lesions, including IDN and LPLK, were identified, and physicians should take note of these. Other imaging methods such as confocal microscopy may aid in the further evaluation of these lesions.